D1.5 Data Management Plan

2025-06-30

Authors: Nikol Yovcheva, Lissa Breugelmans

Funded by
the European Union

D1.5 DMP

Prepared under contract from the European Commission
Grant agreement No. 101180559
EU Horizon Europe Innovation Action
European Research Executive Agency

Project acronym:	**OneSTOP**
Project full title:	**OneBiosecurity Systems and Technology for People, Places and Pathways**
Project duration:	01.01.2025 – 30.06.2028 (42 months)
Project coordinator:	Dr. Quentin Groom, Agentschap Plantentuin Meise (MeiseBG)
Call:	HORIZON-CL6-2024-BIODIV-01-1
Deliverable title:	Data Management Plan
Deliverable №:	D1.5
WP responsible:	WP1
Nature of the deliverable:	DMP — Data Management Plan
Dissemination level:	Public
Licence of use:	CC BY
Lead partner:	PENSOFT
Recommended citation:	Yovcheva, N., Breugelmans, L. (2025). ***Data Management Plan.*** OneSTOP project D1.5.
Due date of deliverable:	Month №6
Actual submission date:	Month №6

Deliverable status:

Version	Status	Date	Author(s)
1.0	Draft	2025 May 29	Nikol Yovcheva, Pensoft Publishers. Lissa Breugelmans, Meise Botanic Garden
1.1	Review	2025 June 18	Lien Reyserhove, Institute for Nature and Forest Research Sofie Meeus, Meise Botanic Garden
1.2	Revised draft checked by the consoritum	2025 June 25	Sent to all partner leads

Table of contents

Key takeaway messages

- D1.5 Data Management Plan provides an overview of the data and other research outputs OneSTOP anticipates generating and reusing, as well as how these will be managed in an open and FAIR (Findable, Accessible, Interoperable, and Reusable, Wilkinson et al. 2018) manner.
- Table 1 provides a summary of the datasets to be generated. Table 2 outlines the data identified for reuse. Table 3 presents the other research outputs that the project will produce.
- All recommendations for ensuring that the data and outputs are made FAIR are summarised in Chapter 9.

Executive summary

To ensure the findability, accessibility, interoperability, and reusability of its research data and other outputs, OneSTOP has developed a Data Management Plan (DMP). The DMP builds on the open science, open data, and open-source principles outlined in the project's Description of Action, as well as input gathered through a consortium-wide consultation. It begins with a summary (Chapter 2) of the datasets OneSTOP expects to generate and reuse, including their types, formats, size, expected timelines, and intended users. Chapter 3 then provides guidance for partners on implementing the FAIR principles. The DMP also identifies the other research outputs OneSTOP will produce and explains how the FAIR principles apply to these (Chapter 4). Additional chapters address administrative aspects of data management, including allocation of resources (Chapter 5), data security (Chapter 6), intellectual property rights (Chapter 7), and ethical considerations (Chapter 8). Chapter 9 summarises all the recommendations provided in the DMP.

Non-technical summary

To make sure OneSTOP's results are of maximum value to stakeholders who can benefit from them, we have created a Data Management Plan. This plan is based on the project's commitment to open science and was developed with input from all project partners. It starts by listing the types of data OneSTOP will create or use, how it will be shared, and who might benefit from it. It also explains how the project will follow the FAIR principles to make the data as useful as possible. The plan also covers other types of results the project will produce, such as tools or reports, and how these will be managed according to the same principles. In addition, it looks at practical topics like who is responsible for managing the data, how it will be kept secure, who owns it, and how ethical standards will be upheld.

List of abbreviations

AMI	Automated Monitoring of Insects
AU	Aarhus University
CC BY	Creative Commons Attribution International Public License
CC BY-NC	Creative Commons Attribution-NonCommercial License
CC0	Creative Commons Public Domain Dedication
CIBIO	BIOPOLIS Association
CU	Coventry University
CyI	Cyprus Institute
DMP	Data Management Plan
DOI	Digital Object Identifier

EASIN	European Alien Species Information Network
EML	Ecological Metadata Language
EU	European Union
EUNIS	European Nature Information System
FAIR	Findable, Accessible, Interoperable, Reusable
GBIF	Global Biodiversity Information Facility
GDPR	General Data Protection Regulation
GF	GreenFormation
GRIIS	Global Register of Introduced and Invasive Species
IAS	Invasive Alien Species
IBOT-CAS	Institute of Botany of the Czech Academy of Sciences
INBO	Research Institute for Nature and Forest
IP	Intellectual Property
LL	Living Lab
LU	Lincoln University
LUKE	Natural Resources Institute Finland
MeiseBG	Meise Botanic Garden
MIT	Massachusetts Institute of Technology
ODMAP	Overview, Data, Model, Assessment and Prediction
PK	Platform Kinetics
SU	Stellenbosch University
UKCEH	UK Centre for Ecology & Hydrology
UFZ	Helmholtz Centre for Environmental Research GmbH
UNIVIE	University of Vienna
UOC	Ovidius University of Constanța
UoE	University of Exeter
WP	Work Package

1. Introduction

OneSTOP's Work Package (WP) 1 is responsible for developing and maintaining the project's Data Management Plan (DMP). This plan outlines the types of data and other research outputs to be generated or reused by OneSTOP and explains how they will be made open and aligned with FAIR principles. To support project members in practical implementation, the DMP will be complemented by a concise Data Management Guidelines document with clear, actionable steps. As a living document, the DMP will be updated regularly to reflect the evolving data of the project, with versions scheduled for delivery at M6 (June 2025), M20 (August 2026) and M40 (April 2028).

The DMP is based on the initial Open Data – Open Source approach outlined in the project's description of action and insights from a detailed data management questionnaire (see Annex 1), which was filled out by partners in April and May 2025[1]. Project members contributed by identifying the datasets and research outputs they will generate and reuse, as well as detailing their institutional data management practices. Additionally, the DMP was distributed to representatives from all project organisations prior to its submission.

2. Data summary

OneSTOP will generate datasets spanning a wide range of formats—including spatial raster files (GeoTIFF), spreadsheets (.xlsx – with a recommendation to convert to .csv to ensure FAIRness where relevant), textual reports (.pdf, .doc), image and audio files (.jpg, .WAV), and code outputs (.py, .RData). Their sizes vary from 2 kb for surveys to 1 TB for high-resolution species distribution models. Data collection methods include field-based tools such as car-mounted cameras, air-DNA samplers and Automated Monitoring of Insects (AMI) traps, as well as modelling and machine learning, stakeholder observations, public surveys and online analysis.

The generated datasets will be made openly accessible, using either a public domain dedication (CC0) or open licences such as CC BY or CC BY-NC, depending on the sensitivity and intended use of the data. Sensitive and personal data, when present, will be aggregated or anonymised in accordance with ethical standards. Partners will use commonly accepted data and metadata standards such as Darwin Core and the Ecological Metadata Language (EML). Where no common standards are used, at the very least, partners will include a set of fields such as Digital Object Identifier (DOI), creators, titles, dates, formats, sizes, licence, applied filters, species names, methods, as well as used protocols and tools, and anonymisation methods. Where necessary, partners will provide additional support for reuse, including README files, software or tools documentation and annotated R Markdown scripts.

Global infrastructures such as Zenodo, the European Alien Species Information Network (EASIN) and the Global Biodiversity Information Facility (GBIF) will be used for data dissemination. The intended user base includes governmental and regulatory bodies, industry and private sector, environmental and conservation groups, research and technology groups, society and the general public. Only six datasets were reported, which

[1] The information provided is subject to change in order to reflect the project's development and arising data needs since the questionnaire was filled out in the very early stage of the project. Updated information, as well as adjustments to the current information, will be provided on an as-needed basis and at the latest during the planned DMP updates before the end of the project.

could potentially contain observation data on species of conservation concern that could lead to their exploitation if site locations are not protected (i.e., rows 5, 7, 8, 16 and 22). To mitigate that risk, exact geographic coordinates will be removed from final datasets and the data will be anonymised (using AI algorithms in the case of CamAlien). Serving as a preliminary overview, Table 1 presents a summary of the datasets that the project expects to generate, whereas Table 2 outlines the data intended for reuse.

Table 1: Summary of the data OneSTOP anticipates generating[2]

No	Dataset name	Contact	Relevant task	Generated via	Size	Format	Sensitive data	Personal data	Publication timeline	Users	Type of access	Licence	Point of access
1	Public Interest in Species of Union Concern (PISUC)	IBOT-CAS	T6.1	Conservation culturomics	~ 2 MB	.xlsx[3]	No	No	M7 July 2025	Research and Technology groups	Open	CC0	Zenodo
2	Stakeholder mapping and baseline surveys for LL members	INBO, MeiseBG, LUKE, CU, UoC, BIOPOLIS -CIBIO	T3.1	Survey	400 KB	.xlsx[4]	Gender, Ethnicity	Organisation, Name	M10 October 2025	OneSTOP, Governmental and regulatory bodies	Open, anonymised	CC BY	Zenodo
3	IAS Distribution Models	BIOPOLIS -CIBIO	T5.1	Modelling, machine learning, statistical analysis	~1 Tb	GeoTIFF , CSV	No	No	M17 May 2026	Research and Technology groups, Industry and private sector	Open	CC0	Zenodo
4	Potential Introduction Risk Maps	BIOPOLIS -CIBIO	T5.1	Spatial mapping, risk assessment	250-500 Mb	GeoTIFF	No	No	M17 May 2026	Research and Technology groups, Industry and private sector	Open	CC0	Zenodo
5	iEcology data	MeiseBG	T2.3	Data aggregation and processing	~5 GB	.csv, .html, .png, .svg	No	Aggregated location	M20 August 2026	Research and Technology groups, Governmental and regulatory bodies	Open, anonymised	CC BY	Zenodo
6	GoogleTrend USA analysis	LU	T2.3	Extraction from web data	5 Mb	.xlsx[5]	No	No	M24 December 2026	OneSTOP	Open	CC0	Lincoln University Data Repository

[2] The information provided in Table 1 is based on an initial data mapping and is subject to change in order to reflect the project's development and arising data needs. Updated information, as well as adjustments to the current information, will be provided on an as-needed basis and at the latest during the planned DMP updates before the end of the project.
[3] Consider converting to .csv to ensure FAIRness if relevant
[4] Consider converting to .csv to ensure FAIRness if relevant
[5] Consider converting to .csv to ensure FAIRness if relevant

No	Dataset name	Contact	Relevant task	Generated via	Size	Format	Sensitive data	Personal data	Publication timeline	Users	Type of access	Licence	Point of access
7	Automated IAS reports	UKCEH	T5.6	TBD	1-5GB	.html, .pdf	No	No	M24 December 2026	TBD	Open	TBD	TBD
8	Illumina sequence data (raw)	UKCEH	T2.2	Illumina Sequencing	20GB per file, 100GB total	FASTQ	No	No	M24 December 2026	TBD	Open	TBD	TBD
9	ASV Tables	UKCEH	T2.2	Bioinformatics processed raw data	1MB	.csv	No	No	M24 December 2024	TBD	Open	TBD	TBD
10	Survey results	GF	T6.2	Survey	~ 5 MB	.xlsx[6], .pdf, .csv	No	No	M25 January 2027	Research and Technology groups, Governmental and regulatory bodies	Open	CC0 or CC BY	Zenodo
11	Potential current and future distribution maps of terrestrial native species in Europe threatened by invasive alien species	UNIVIE	T5.4	Modelling	TBD	tiff; RData	No	No	M26 February 2027	Research and Technology groups, Industry and private sector	Open	CC-BY-NC	Zenodo & GitHub
12	Horizon Scanning Output	MeiseBG	T5.3	A semi/automated workflow	1 MB	.csv	No	No	M32 August 2027	Research and Technology groups, Governmental and regulatory bodies	TBD	CC0	TBD

[6] Consider converting to .csv to ensure FAIRness if relevant

No	Dataset name	Contact	Relevant task	Generated via	Size	Format	Sensitive data	Personal data	Publication timeline	Users	Type of access	Licence	Point of access
13	Priority Species Database	CIBIO-BIOPOLIS	T5.5	Modelling, optimisation and prioritisation, Multi-criteria analysis	~1 Tb	Spatial ORDBMS, GeoTIFF, CSV	No	No	M33 September 2027	Research and Technology groups, Industry and private sector	Open	CC0	Zenodo
14	Results from a multi-stakeholder scenario workshop in LLs	INBO, MeiseBG, LUKE, CU, UoC, BIOPOLIS-CIBIO	T3.3	Questionnaire	1 Mb	.xlsx[7]	No	No	M33 September 2027	Research and Technology groups, LL stakeholders	Open, anonymised	TBD	TBD
15	Stakeholder feedback on the experimentation in LLs	INBO, MeiseBG, LUKE, CU, UoC, BIOPOLIS-CIBIO	T3.2	Survey	1 Mb	TBD	TBD	Yes	M36 December 2027	Research and Technology groups, LL stakeholders, Industry and private sector	Open	TBD	TBD
16	CamAlien	AU	T2.1	Vehicle-mounted camera	20 TB	.jpg	No	Potentially	After M37 After January 2028	Governmental and regulatory bodies, industry and private sector	Open	CC BY	erda.au.dk
17	AMI	AU	T2.1	Insect camera trap	10 TB	.jpg	No	Yes	After M37 After January 2028	Industry and private sector	Open	CC BY	erda.au.dk
18	Dataset UOC #1	UOC	T2.4	Observation data	TBD	.xlsx[8]	No	Yes (IP address)	M38 February 2028	Research and technology groups, society and the general public	Open	CC0	iNaturalist, GBIF

[7] Consider converting to .csv to ensure FAIRness if relevant
[8] Consider converting to .csv to ensure FAIRness if relevant

No	Dataset name	Contact	Relevant task	Generated via	Size	Format	Sensitive data	Personal data	Publication timeline	Users	Type of access	Licence	Point of access
19	Dataset UOC #2	UOC	T6.3	Interviews	TBD	.xlsx[9]	No	Yes (name, surname, email)	M38 February 2028	Researchers, policy makers, managers	Only for OneSTOP	N/A	N/A
20	Maps of IAS potential distribution	UFZ	T5.2	Modelling	~ 50 kB per map	TIFF	No	No	M39 March 2028	Research and Technology groups, Governmental and regulatory bodies	Open	CC BY	Zenodo
21	Prioritisation maps for IAS management in Europe	UNIVIE	T5.4	Modelling	TBD	tiff, RData	No	No	M39 March 2028	Research and Technology groups, industry and private sector	Open	CC-BY-NC	Zenodo & GitHub
22	Automated EICAT assessments	UKCEH	T4.4	TBD	1-5GB	.html, .pdf	No	No	M40 April 2028	TBD	Open	TBD	TBD
23	LL evaluation and optimisation	INBO, MeiseBG, LUKE, CU, UoC, BIOPOLIS-CIBIO	T3.4	Survey, LL focus group meetings, Observation	~ 1-10 MB	.doc, .pdf, .jpeg, .WAV	TBD	Yes	M40 April 2028	Governmental and regulatory bodies, industry and private sector, environmental and conservation groups, research and technology groups, society and the general public	Open, anonymised	CC BY	Zenodo

[9] Consider converting to .csv to ensure FAIRness if relevant

No	Dataset name	Contact	Relevant task	Generated via	Size	Format	Sensitive data	Personal data	Publication timeline	Users	Type of access	Licence	Point of access
24	Sentinel Gardens Coventry	CU	T2.4	iNaturalist, Observation	~ 1-10 MB	.pdf, .doc	TBD	Yes	M40 April 2028	Governmental and regulatory bodies, industry and private sector, environmental and conservation groups, research and technology groups, society and the general public	Open	CC BY	GBIF; Zenodo
25	GRIIS Europe	INBO	T4.2	Data aggregation	10.000 records	DwC-A (csv + xml)	No	No	TBD	Research and Technology groups, Governmental and regulatory bodies	Open	CC0	GBIF
26	Multiple thematic checklists of IAS within Europe	INBO	T4.2	Expert aggregation and validation	TBD	.csv	No	No	TBD	Research and Technology groups, Governmental and regulatory bodies	Open	CC0	GBIF
27	Comparative Analysis of Invasive Species Demographics	UoE	T5.6	Database search and primary literature	50 Mb	.xlsx[10]	No	No	TBD	Research and Technology groups	Open	Public domain	Zenodo

In addition to the data intended for reuse in Table 2, OneSTOP partners will repurpose data generated within the project itself to support further task development. For example, IAS observations collected in WP2—using air-DNA, computer vision, iEcology, and citizen science—will be published to GBIF in WP4 as distinct datasets, each assigned a DOI and described using EML. National checklist data will also be made available via GBIF, and subsequently shared with other platforms, including the Global Register of Introduced and Invasive Species (GRIIS) or the Global Invasive Species Database.

[10] Consider converting to .csv to ensure FAIRness if relevant

Table 2: Summary of the data OneSTOP anticipates reusing[11]

No	Dataset name	Contact	Relevant task	Size	Format	Sensitive data	Personal data	Access	Origin	Ownership	Licence
1	Comparative Analysis of Invasive Species Demographics	UoE	T5.6	50 Mb	.xlsx	No	No	Open	Primary literature	Public	Open Access Licences
2	GBIF occurrence data from IAS in Europe	INBO	T4.3	> 1M records	.csv, .xml	No	No	Open	GBIF	Public	CC0
3	National lists of introduced and invasive species within Europe	INBO	T4.2	> 100K records	.csv, .xml	No	No	Open	GBIF	Public	CC0
4	GBIF Occurrences data	NNIT	T4.3	> 1M records	Darwin Core Archive	No	No	Open	GBIF	Various	CC licences
5	CHELSA Bioclimatic indices and future projections	CIBIO-BIO POLIS	T5.1	~0.5Tb	GeoTIFF, NetCDF	No	No	Open	CHELSA (https://chelsa-climate.org/)	Chelsa Climate	CC BY
6	Land Use: current and future projections	CIBIO-BIO POLIS	T5.1	~0.5Tb	GeoTIFF, NetCDF	No	No	Open	https://doi.org/10.25584/data.2020-07.1357/1644253 https://doi.org/10.5281/zenodo.3713432 https://doi.org/10.5281/zenodo.3713378	Pacific Northwest National Laboratory 2; PNNL	CC BY

[11] The information provided in Table 2 is based on an initial data mapping and is subject to change in order to reflect the project's development and arising data needs. Updated information, as well as adjustments to the current information, will be provided on an as-needed basis and at the latest during the planned DMP updates before the end of the project.

No	Dataset name	Contact	Relevant task	Size	Format	Sensitive data	Personal data	Access	Origin	Ownership	Licence
7	Bioclimatic data	UFZ	T5.2	~500 MB	.nc	No	No	Open	CHELSA (https://chelsa-climate.org/)	Chelsa Climate	TBD
8	Species occurrence data	UFZ	T5.2	TBD	.csv or .RData	TBD	No	Open	GBIF	GBIF	CC BY-NC
9	EEA Reference grid	UFZ	T5.2	TBD	GDB	No	No	Open	SDI - geospatial data catalogue - EEA (https://sdi.eea.europa.eu/)	European Environment Agency	CC BY
10	Mined social media data	MeiseBG	T2.3	~ 3 GB	.csv	No	Coordinates, User names	Restricted	Wikidata, Youtube, Flickr, Bluesky, Mastodon, iNaturalist, GBIF, eBay	Respective platforms	According to respective platforms
11	Occurrence records (native, invasive alien species)	UNIVIE	T5.4	TBD	.csv; .txt	No	No	Open	GBIF	Various	CC0, CC BY, CC BY-NC
12	Climate Data	UNIVIE	T5.4	285.09 KiB	.tiff	No	No	Restricted	CHELSA	CHELSA	CC BY
13	Protected Areas	UNIVIE	T5.4	1.3GB	.fgbd	No	No	Open	WDPA	UNEP-WCMC	CC BY-NC
14	Land-cover data	UNIVIE	T5.4	5GB	GeoTIFF	No	No	TBD	CORINE land cover	European Environment Agency EEA	TBD

3. FAIR data

This chapter provides recommendations on applying the FAIR data principles to OneSTOP's data. In the context of OneSTOP's DMP, the term research data covers both primary (raw or input data) and processed (output) data that underpin the project's results.

Before diving into the individual FAIR principles, it is important to emphasise one overarching practice that supports all aspects of FAIRness—ensuring that datasets are uploaded to trusted open access repositories. This foundational step promotes long-term accessibility, enhances findability, enables interoperability through standard formats and metadata, and facilitates reuse by making data openly available under clear licensing terms. Encouraging all partners to follow this practice is essential for achieving meaningful FAIR alignment.

- **Recommendation #1. Trusted repositories**. Upload datasets to trusted repositories[12] and infrastructures such as Zenodo or GBIF, which assign a DOI and allow easy download via a simple https link, without the need for special tools for access. This allows anyone with the link to retrieve the data through a standard web browser, provided the data are open. In terms of a general-purpose trusted repository, the majority of OneSTOP partners expressed a preference for Zenodo as a suitable option, which they are familiar with and have previously explored appropriate arrangements regarding deposition. Thus, when not using domain-specific repositories, OneSTOP partners are encouraged to deposit their datasets to Zenodo as the project can verify that it covers all the requirements of FAIR data, most importantly:
    - A DOI is issued to every published dataset on Zenodo;
    - Metadata of each record are indexed and searchable directly in Zenodo's search engine immediately after publishing;
    - Metadata of each dataset are sent to DataCite servers during DOI registration and indexed there;
    - Metadata for individual records as well as record collections are harvestable using the standard, open, free and universal OAI-PMH protocol by the record identifier and the collection name;
    - Metadata are publicly accessible and in the public domain. No authorisation is ever required to retrieve it.
    - Data and metadata will be retained for the lifetime of the repository, ensuring that the metadata will be accessible, even when the data are no longer available;
    - Metadata are stored in high-availability database servers at CERN, which are separate from the data itself.

[12] Trusted repositories are either certified repositories (e.g., CoreTrustSeal) and/or disciplinary/domain repositories that are commonly used/endorsed by research communities (e.g., ELIXIR deposition databases). General-purpose (e.g., Zenodo) and institutional repositories could also be acceptable if they share these properties: Mechanisms to ensure integrity and authenticity of contents; Clear information about their policies/services; Broad, ideally open access to content (consistent with legal and ethical constraints); Assigned PIDs; Detailed, machine-readable and standardised metadata; Mid- and long-term preservation of contents, expert curation, quality assurance; National and/or international security criteria.

3.1. Making data findable

The first step in maximising data value is ensuring it can be easily found by both humans and machines. This requires assigning it globally unique identifiers, providing rich metadata, and registering both data and metadata in searchable resources.

- **Recommendation #2 - Identifiers**. Publish data with globally unique and persistent identifiers. You can find a list of services that supply such identifiers here.
- **Recommendation #3 - Metadata**. Provide machine-readable, detailed metadata adhering to standardised metadata formats. Include descriptive information about the context, quality and condition of the data, using keywords which enable computers to automatically sort and prioritise data.
    - **Recommendation #3.1**. OneSTOP has identified a number of suitable standards that can be used based on each partner's preference and specific data needs:
        - Ecological Metadata Language (EML): includes modules for the spatial, temporal, taxonomic, and thematic extent of data, as well as research methods and protocols.
        - DataCite Metadata: provides a standardised framework for data across disciplines.
    - **Recommendation #3.2**. The metadata should explicitly reference the dataset's globally unique and persistent identifier.
    - **Recommendation #3.3**. If no specific metadata standard is applied, at the very least, include the following fields:
        - Creator(s)
        - Creation date
        - Modified date
        - Publication date
        - Title
        - Identifier
        - Dataset description
        - Dataset language
        - Input data/parameters/filters/protocols used to generate the dataset
        - Ontology-aligned keywords and meta-tags
        - Version
        - Size
        - Format
        - Project acronym (OneSTOP) and number (101180559)
        - Licence of use
        - Embargo period
        - Publisher

- **Recommendation #4 - Indexing**. To ensure your research (meta)data are findable online, choose data repositories that register, index, and make your (meta)data discoverable.

3.2. Making data accessible

Once users find the data, they must understand how to access it. This means minimising access barriers by storing data in trusted repositories, making (meta)data retrievable via a standardised communication protocol using its identifier, and ensuring that metadata remains accessible even if the data is no longer available.

- **Recommendation #5 - Access protocols**. Research data should be easily accessible online using standard, free and open protocols, rather than relying on specialised or proprietary tools or communication methods. If there are conditions and/or you are not using a free and open protocol for accessing the data, ensure that the requirements are clearly defined, allowing automated systems to handle or notify users.

OneSTOP will follow the principle of data being “as open as possible as closed as necessary”, and restricted access may be implemented for specific datasets, if deemed necessary. Table 1 provides an overview of the datasets designated as restricted or closed (if any), along with the corresponding reasons. Partners will review the finalised datasets to assess the possibility of sharing them in aggregated and/or anonymised form to safeguard sensitive and/or personal information. Table 1 also includes the proposed embargo periods (if any) for certain datasets.

3.3. Making data interoperable

After users find and access OneSTOP’s data, they should ideally discover that the datasets are formatted and presented in a way that enables their interoperability with applications or workflows for analysis, storage, and processing. OneSTOP will make sure its datasets can be integrated with other datasets, using broadly accepted languages for knowledge representation, standard vocabularies, and including references to other relevant metadata.

- **Recommendation #6 - Language**. Use a formal, accessible, shared, and broadly applicable language for knowledge representation, such as the Resource Description Framework, Web Ontology Language, DARPA Agent Markup Language or the Javascript Object Notation for Linking Data.
- **Recommendation #7 - Standards and vocabularies**. Use data standards and vocabularies such as Darwin Core and the European nature information system (EUNIS) that are documented and accessible to anyone using the dataset.
- **Recommendation #8 - References**. Include qualified references to and properly cite other (meta)data by specifying if one dataset builds on another, if additional datasets are needed to complete the data, or if complementary information is stored in a different dataset.

3.4. Making data reusable

As outlined in Chapter 3.2, OneSTOP’s data will be made freely available in the public domain under licences that allow reuse, facilitating the broadest possible use by third parties, even beyond the project’s completion. To help users assess the relevance of the data for their specific context, the datasets will be published with standardised, rich metadata (as detailed in Chapter 3.1). This metadata will provide key information on the data’s provenance, scope, limitations, generation process, technical specifications, related species, and more. To prevent ambiguity that might impede data reuse due to unclear licensing terms, OneSTOP’s metadata will clearly state licensing conditions in a format that is both human- and machine-readable. Where appropriate, supporting documentation necessary for analysing, validating, and reusing the data will also be provided, including README files, software or tool documentation, and annotated R Markdown scripts.

- **Recommendation #9 - Clear licensing**. Release data with a clear data usage licence.
- **Recommendation #10 - Supporting documentation**. Information should be provided about any research output or any other tools and instruments needed to re-use or validate the data. Describe data not only with standardised metadata

supporting discovery but also provide readme files or documentation needed for data analysis, validation and reuse with information on data provenance, scope, limitations, generation process, etc., such as the following:

- Data origin;
- Data provider;
- Date of data retrieval;
- Licence of use;
- A URL or DOI for the original data;
- Description of the input data, including its format, structure, and content;
- Steps taken to clean and preprocess the data to ensure data quality and integrity.

3.5. Making data open

While aligning with the FAIR principles, it is important to note that a FAIR dataset does not necessarily mean an open one, and open does not automatically mean FAIR. A dataset can be FAIR while still having access restrictions, and conversely, an open dataset may lack the metadata, structure, or standards compliance needed to be truly FAIR.

In line with Article 17 of the Horizon Europe Grant Agreement, and the principle of "as open as possible, as closed as necessary," datasets should be made open by default, unless specific reasons prevent this. Open access may be restricted if it would compromise the beneficiary's legitimate interests, such as commercial exploitation, or if it conflicts with other regulations (e.g., GDPR) or obligations, including EU competitive interests or terms of the Grant Agreement. In such cases, any decision to withhold open access must be clearly justified in the project's DMP.

- **Recommendation #11 - Open data rules**. According to Article 17 of the Grant Agreement, OneSTOP partners should make their data open access[13] by:
  - **Recommendation #11.1**. Uploading the data to an open access trusted data repository under the latest available version of the Creative Commons Attribution International Public License (CC BY) or Creative Commons Public Domain Dedication (CC0) or a licence with equivalent rights. Partners should preferably use a thematic repository like GBIF. If no such repository exists, they could use a generic repository such as Zenodo or a trusted institutional repository.
  - **Recommendation #11.2**. The data underlying scientific publications should be deposited in a repository at the time of paper publication and made openly accessible as soon as possible, but no later than 12 months after the paper's publication. Other research data should be deposited and made open at the earliest opportunity.
  - **Recommendation #11.3**. Information should be provided via the repository about any research output or any other tools and instruments needed to re-use or validate the data.
  - **Recommendation #11.4**. The metadata of deposited data must put in the public domain under a CC0 licence waiver or another equivalent waiver (to

[13] Unless providing open access would, in particular, 1) be against the beneficiary's legitimate interests, including regarding commercial exploitation, or 2) be contrary to any other constraints, in particular the EU competitive interests or the beneficiary's obligations. If open access is not provided (to some or all data), partners must inform Pensoft in order to include a justification in the DMP.

the extent legitimate interests or constraints are safeguarded). Datasets may become unavailable over time due to maintenance costs, but metadata should persist, even if the data are no longer accessible. This ensures valuable information remains searchable and useful for research, including in replication studies.[14]

4. Other research outputs

Besides ensuring FAIR data management, projects should also consider the management of the other research outputs that they may generate or reuse. In OneSTOP's case, such outputs include workflows, species lists, DNA samples, Shiny apps and hardware used in WP2. Table 3 provides an overview of the anticipated research outputs for generation.

[14] In cases where data are closed but there are no compelling reasons that the related metadata should be, it is recommended that open access is provided to the metadata of the data, with CC0 or equivalent, while the dataset itself remains closed.

Table 3: Summary of the other research outputs OneSTOP anticipates generating[15]

No	Output name	Contact	Relevant task	Type of output	Delivery date	Access	Licence	Users	Point of access
1	OneSTOP alert tool	NNIT	T4.3	Software, Early Warning System	After M12 December 2025	Open	MIT licence	Research and technology groups, Industry and private sector	GitHub
2	Species list repository	CIBIO-BIO POLIS	T5.1, T5.5	A harmonised species list intended for modelling and prioritisation	M17 May 2026	Open	CC0	Industry and private sector, environmental and conservation groups, research and technology groups	GitHub
3	Conservation culturomics workflows	IBOT-CAS	T6.1	R scripts	M18 June 2026	Open	MIT for software, CC0 for data	Research and technology groups	GitHub
4	iEcology workflow	MeiseBG	T2.3	R scripts	M20 August 2026	Open	MIT licence for software, CC0 for data	Research and technology groups, Governmental and regulatory bodies	GitHub, Zenodo
5	eDNA samples	LUKE	T2.2	DNA samples collected by aerial spore/pollen trap	M24 December 2026	Physical samples, no direct access	N/A	Research and technology groups	N/A
6	eDNA workflow	LUKE	T2.2	Report/best practices	M24 December 2026	Open	TBD	LL stakeholders, Industry and private sector	TBD

[15] The information provided in Table 3 is based on an initial data mapping and is subject to change in order to reflect the project's development and arising data needs. Updated information, as well as adjustments to the current information, will be provided on an as-needed basis and at the latest during the planned DMP updates before the end of the project.

No	Output name	Contact	Relevant task	Type of output	Delivery date	Access	Licence	Users	Point of access
7	Horizon scanning	MeiseBG	T5.3	Script	M32 August 2027	Open	MIT licence for software, CC0 for data	Research and technology groups	Zenodo
8	Instrumentation	Platform Kinetics	T2.2	Hardware	M36 December 2027	Open	N/A	Living Labs	N/A
9	SDM workflow	UFZ	T5.2	R scripts	M39 March 2028	Open	CC BY	Research and technology groups	GitHub
10	Prioritisation maps for IAS management in Europe	UNIVIE	T5.4	Workflow	M39 March 2028	Open	TBD	Research and technology groups	GitHub, Zenodo
11	Policy brief	SUN & UOC	T6.4	Report	M41 May 2028	Open	CC BY	Governmental and regulatory bodies	TBD
12	Invasive Species Living Lab Guidelines	CU	T3.4	Guideline	M42 June 2028	Open	CC BY	Governmental and regulatory bodies, industry and private sector, environmental and conservation groups, research and technology groups, society and the general public	Zenodo
13	Shiny Apps for modelling invasive species	UoE	T1.4	Shiny Web-app	TBD	Open	TBD	Society and the general public	TBD
14	Reproducible publication workflows for IAS detection data and IAS checklists	INBO	T4.1, T4.2	Software (R code)	TBD	Open	MIT Licence	Industry and private sector, research and technology groups	GitHub

4.1. Outputs management

As outlined in Table 3, OneSTOP's other research outputs will be made open and accessible, with most licensed for reuse under CC BY or MIT licences. Where relevant and applicable (e.g., excluding physical samples), these will be deposited in open access repositories such as Zenodo and GitHub to ensure broad availability and reuse.

Partners will strive to assign globally unique and persistent identifiers to their other research outputs which will be described with rich machine-readable metadata. Open source software produced during the project will be managed on GitHub, and releases will be archived to Zenodo to provide versioning and a DOI so they are uniquely referenceable. The metadata will cover essential descriptive, operational and structural elements, such as sample name, collection date, location, collector information, and usage rights. Standard metadata fields will include DOI, creators, titles, dates, file sizes, formats, descriptions, and licensing terms.

To ensure interoperability and reusability, scripts will be made openly available through a well-structured GitHub repository. Software code will be managed and versioned in the repository and will be programmed in open-source languages. Bugs and feature requests will be tracked using an issue tracker linked to numbered versions. Workflows will adhere to FAIR principles and employ standardised, widely accepted formats such as CSV for data and R scripts for extraction and analysis. Comprehensive documentation will be provided through README files and RMarkdown. Furthermore, to enhance the standardisation and documentation of model protocols, OneSTOP partners will follow the ODMAP (Overview, Data, Model, Assessment and Prediction) standard protocol for reporting species distribution models.

- **Recommendation #12 - Management of other research outputs**. OneSTOP partners are encouraged to manage their other research outputs in accordance with the FAIR principles described in Chapter 3.:
    - **Recommendation #12.1.** Strive to deposit your research outputs described with rich metadata in trusted open access repositories which assign them globally unique and persistent identifiers and offer search engines and indexing.
    - **Recommendation #12.2.** Use standard formats, vocabularies and ontologies and accompany your outputs with a separate human-readable description of the output, where needed.
    - **Recommendation #12.3.** Strive to make your outputs accessible under open licences, such as CC BY, CC0 and the MIT Licence. When using the MIT Licence, consider using the full text and disclaimer of the licence:

      MIT License

      Copyright (c) [year] [full name]

      Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

In addition to the outputs listed in Table 3, OneSTOP will generate 36 public deliverables and 23 milestones. Subject to author approval, these will be made accessible through the project's online library. Furthermore, OneSTOP will establish a dedicated, citable collection in the open access journal Research Ideas and Outcomes. This collection will serve as a central hub for publications, data papers, policy briefs and project deliverables—each with a permanent DOI to ensure the project's collective knowledge remains accessible and citable beyond its duration. Once deliverables have undergone external review and received official approval, OneSTOP partners will have the option of publishing them as preprints within this collection. Findings will also be disseminated through gold open access publications[16], covering the different aspects of the project.

- **Recommendation #13 - Scientific publications**. OneSTOP partners should adhere to the following open access rules for publications:
    - **Recommendation #13.1.** Deposit in a trusted repository, at the latest at the time of publication, a machine-readable electronic version of the published version or the final peer-reviewed manuscript accepted for publication;
    - **Recommendation #13.2.** Provide immediate open access via the repository, under the latest available version of the Creative Commons Attribution International Public License (CC BY) or a licence with equivalent rights;
        - For monographs and other long-text formats, the licence may exclude commercial uses and derivative works (e.g., CC BY-NC, CC BY-ND);
    - **Recommendation #13.3.** Give information via the repository about any research output or any other tools and instruments needed to validate the conclusions of the scientific publication;
    - **Recommendation #13.4.** Include a Data Availability Statement (even when there is no associated data) with information on where data can be accessed, preferably not referring readers to contact authors in order to access the data.

OneSTOP will not only generate new research outputs but also reuse existing ones. Key external resources to be reused include the codebase of the GBIF alert tool, which will help develop OneSTOP's early alert system under Task 4.3. The alien species risk modelling and mapping framework, from the TrIAS project, will be reused in Task 5.1, with the code openly available on GitHub under the MIT licence. OneSTOP will also make use of specialised monitoring tools, including two CamAlien camera systems for tracking invasive alien plants,

[16] Only publication fees in full open access venues for peer-reviewed scientific publications are eligible for reimbursement. You can check the eligibility of the journal via this website.

three AMI traps for insect monitoring and eDNA filters, which will be tested in the Living Labs. Furthermore, modelling work from Task 5.1 will be reused in Task 5.4 to support the development of distribution maps.

4.2. Software development recommendations

Software produced by the project should be open, well-documented and developed collaboratively using best practices. In order to achieve this, software will be managed in a GitHub repository which contains all files, discussions and version history related to a single software package or analysis. Bugs and feature requests will be tracked using an issue tracker linked to numbered versions. Below we present a list of recommendations related to the development of software within OneSTOP. We refer to the B-Cubed software development guide for more details on the implementation of these recommendations, and further guidelines on best practices.

- **Recommendation #14 - Software development guidelines:**
  - **Recommendation #14.1.** Code should be developed in a public GitHub repository as part of the OneSTOP organisation. The repository should be set up in a structured way and contain a README.md, .gitignore, CITATION.cff, CODE_OF_CONDUCT.md and LICENCE file with mention of the copyright holder (recommended: MIT). Code that already exists (locally), should be added to GitHub following these guidelines. In case code has already been developed in a private GitHub repository, the repository should be transferred to the OneSTOP organisation (rights to transfer can be requested with one of the organisation owners).
  - **Recommendation #14.2.** Code contributions should follow the GitHub flow. The repository's main branch should contain the software code in a state that can be installed without issue. Changes by collaborators should be made in separate branches and should be reviewed by collaborators through a pull request before merging into the main branch.
  - **Recommendation #14.3.** Developed code should be properly documented. Some guidelines for R function documentation can be found here. The repository should also contain a README file which explains the rationale and scope of the software, the instructions for installation, and a minimal example of the software workflow. Additional examples can be written out in the form of tutorials, which can be included on the project website.
  - **Recommendation #14.4.** Semantic versioning should be used throughout the software development process. Major and minor versions of the software should have an associated GitHub release, and should be published automatically to Zenodo. Version changes should be communicated and explained in a changelog in the form of a NEWS.md file which is added to the repository and kept up-to-date.
  - **Recommendation #14.5.** Code should be written in open-source languages, and best practices for coding style should be followed (as detailed in the B-Cubed software development guide).
  - **Recommendation #14.6.** Software developed within OneSTOP should be subjected to a review process. The maintainer of the software should reach out to OneSTOP partners, or to colleagues from their organisation, to execute a voluntary software review of their developed code, at least once and before the deliverable pertaining to software is submitted.

5. Allocation of resources

FAIR data and output management involves certain costs, such as article processing charges for publishing in gold open access journals, fees for depositing data and code in open access repositories and costs for the use of physical outputs such as monitoring equipment. To address these needs, OneSTOP has allocated appropriate resources to its partners who are expected to use their funding responsibly, with a strong emphasis on supporting open access practices. Dedicated resources have also been provided to the WP1 leaders to ensure the Data Management Plan reflects the evolving data needs of the project.

6. Data security

As detailed in Chapter 3, once OneSTOP's open access datasets and outputs are fully compiled, they will be stored in trusted repositories such as Zenodo, GBIF and GitHub to ensure their long-term preservation and curation. Before that, project partners will securely store them on their respective servers, which are equipped with fault-tolerant systems and benefit from regular and/or automated backup protocols. Access to these servers is tightly controlled through security measures including encryption, passwords, two-factor authentication, VPN connections, and/or firewall protection. Partners are also informed of who the Data Protection Officer is within their institution to ensure compliance with data protection regulations. All personal and un-anonymised digital data will be accessible only to designated OneSTOP team members involved in these tasks. Before digitalisation, hard copies of participants' informed consent forms will be securely stored in locked filing cabinets, and all personal data will be destroyed at the conclusion of OneSTOP.

7. Intellectual property management

Intellectual Property (IP) management within OneSTOP will be conducted in accordance with the Consortium Agreement (chapters 8, 9 and 10) and the Grant Agreement (article 16). Their principles are summarised in this chapter.

7.1. Ownership

'Results' means any tangible or intangible effect of the project, such as data, know-how or information, whatever its form or nature, whether or not it can be protected, as well as any rights attached to it, including intellectual property rights.

Each partner retains ownership of their pre-existing results (background IP). OneSTOP members will provide each other with access to the background IP needed to implement the project's tasks. Use rights will be granted under a non-exclusive, royalty-free licence for project purposes.

Results developed as part of OneSTOP (foreground IP) will be owned by the party(ies) that generate them. If developed collaboratively, the IP will be jointly owned, governed by the Grant Agreement (Article 16.4) and the Consortium Agreement (Chapter 8). Unless otherwise agreed, joint owners may use the results for non-commercial research and teaching without charge or prior consent, and may also exploit or licence them to third parties (non-exclusively and without sub-licensing rights) with at least 45 days' notice and

fair compensation to co-owners. Joint owners must agree in advance on protection measures and cost-sharing.

Confidential information remains the property of the disclosing party. No rights are transferred or implied, and recipients must not apply for IP protection on such information. Confidentiality obligations apply for 5 years after project end, with strict non-disclosure, use limitation, and protection requirements.

The granting authority does not claim ownership of results generated within the project. It is granted non-exclusive, irrevocable, royalty-free rights to use non-sensitive project outputs (e.g., deliverables, publications, audio-visual material) for policy, publicity, and communication purposes, both during and after the project. If materials or documents are subject to moral rights or third-party rights (including intellectual property rights or rights of natural persons on their image and voice), the beneficiaries must ensure that they comply with their obligations under the Grant Agreement (in particular, by obtaining the necessary licences and authorisations from the rights holders concerned).

Tables 2 and 3 list the partner institutions designated as primary contacts for inquiries regarding OneSTOP's results, including questions related to ownership.

7.2. Protection

Foreground IP with potential commercial value will be assessed by the originating partner(s) for appropriate protection measures, such as copyright, database rights, or patent applications. These decisions will be made in consultation with the project coordinator and institutional IP officers, ensuring alignment with the project's dissemination and publication timelines. In the event of disputes arising from IP, decisions are made by the General Assembly, as outlined in the Consortium Agreement.

During the project's initial stages (April and May 2025), partners were asked to provide a preliminary plan for the licences under which their results would be made available and to indicate any anticipated need for IP protection. No partners expressed a need for special protection measures at that stage. Tables 2 and 3 provide an overview of the expected data and other research outputs, including their licensing terms. The selected licence will guide how the data/output can be used, shared, and attributed, while also determining the scope of ownership rights and any necessary protection measures such as copyright or patents.

This information will be updated in M20 (August 2026) and M40 (April 2028), when partners will be consulted again on licences and necessary protection strategies. Sensitive or confidential data will be clearly marked and managed according to established data protection and security protocols. All outputs will include appropriate attribution, including funding acknowledgements and contributor credit.

8. Ethics

All activities developed in OneSTOP must comply with ethical principles and relevant national, EU and international legislation, such as the Charter of Fundamental Rights of the European Union, the European Code of Conduct for Research Integrity and the European Convention on Human Rights. In its Grant Agreement, OneSTOP conducted an Ethics self-assessment, covering key areas of concern, such as the involvement of human

participants, collaborations with non-EU countries, the use of data on protected species and habitats, and the application of artificial intelligence.

Involvement of human participants in the project includes questionnaires, Living Labs, workshops, interviews, and citizen science initiatives. It will follow GDPR and informed consent will be secured for all human participants. Children and individuals unable to provide consent will participate only under parental or legal supervision, for instance, in activities like bioblitzes in local parks or botanic gardens. Non-EU partners like South Africa, the United Kingdom, and New Zealand will be engaged on equal terms, ensuring fair benefit-sharing and local compliance. Research on fauna and flora will avoid direct interference with endangered species and will take precautions not to disclose sensitive location data. OneSTOP uses computer vision and large language models to support invasive species detection and information synthesis, with AI designed to enhance, not replace, human decision-making; thus, automated decision-making will not be used. Specific safeguards include removing images of people from camera data and grounding generated content in verified sources to reduce the risk of false statements.

To ensure ongoing ethical compliance, OneSTOP established an Ethics Committee (Alex Franklin, Coventry University; Agnes Zolyomi, GreenFormation; and Quentin Groom, Meise Botanic Garden) reporting to the Executive Board. The committee is responsible for producing a comprehensive ethics report (D1.8, Month 18, June 2026), which will review informed consent processes, participation procedures, and the handling of personal data. To support ethical engagement with human participants, the committee compiled a OneSTOP Participant Information Sheet and Informed Consent Sheet for use across all partner activities involving human volunteers as participants. At the outset of the project, partners also received a guidance document titled *Guidelines for Partners Implementing Gender Dimension and Promotion of Equality, Diversity and Inclusion in the OneSTOP Project*. In March 2025, an initial ethics survey was distributed to identify which partners required additional support from Coventry University regarding formal ethics clearance. Partners were reminded of their responsibility to secure ethics approval where applicable to their research. The full details on Ethics in OneSTOP will be available in D1.8.

9. All recommendations

- **Recommendation #1. Trusted repositories**. Upload datasets to trusted repositories[17] and infrastructures such as Zenodo or GBIF, which assign a DOI and allow easy download via a simple https link, without the need for special tools for access. This allows anyone with the link to retrieve the data through a standard web browser, provided the data are open. In terms of a general-purpose trusted repository, the majority of OneSTOP partners expressed a preference for Zenodo as a suitable option, which they are familiar with and have previously explored appropriate arrangements regarding deposition. Thus, when not using domain-specific repositories, OneSTOP partners are encouraged to deposit their

[17] Trusted repositories are either certified repositories (e.g., CoreTrustSeal) and/or disciplinary/domain repositories that are commonly used/endorsed by research communities (e.g., ELIXIR deposition databases). General-purpose (e.g., Zenodo) and institutional repositories could also be acceptable if they share these properties: Mechanisms to ensure integrity and authenticity of contents; Clear information about their policies/services; Broad, ideally open access to content (consistent with legal and ethical constraints); Assigned PIDs; Detailed, machine-readable and standardised metadata; Mid- and long-term preservation of contents, expert curation, quality assurance; National and/or international security criteria.

datasets to Zenodo as the project can verify that it covers all the requirements of FAIR data, most importantly:
    - A DOI is issued to every published dataset on Zenodo;
    - Metadata of each record are indexed and searchable directly in Zenodo's search engine immediately after publishing;
    - Metadata of each dataset are sent to DataCite servers during DOI registration and indexed there;
    - Metadata for individual records as well as record collections are harvestable using the standard, open, free and universal OAI-PMH protocol by the record identifier and the collection name;
    - Metadata are publicly accessible and in the public domain. No authorisation is ever required to retrieve it.
    - Data and metadata will be retained for the lifetime of the repository, ensuring that the metadata will be accessible, even when the data are no longer available;
    - Metadata are stored in high-availability database servers at CERN, which are separate from the data itself.
- **Recommendation #2 - Identifiers**. Publish data with globally unique and persistent identifiers. You can find a list of services that supply such identifiers here.
- **Recommendation #3 - Metadata**. Provide machine-readable, detailed metadata adhering to standardised metadata formats. Include descriptive information about the context, quality and condition of the data, using keywords which enable computers to automatically sort and prioritise data.
    - **Recommendation #3.1**. OneSTOP has identified a number of suitable standards that can be used based on each partner's preference and specific data needs:
        - Ecological Metadata Language (EML): includes modules for the spatial, temporal, taxonomic, and thematic extent of data, as well as research methods and protocols.
        - DataCite Metadata: provides a standardised framework for data across disciplines.
    - **Recommendation #3.2**. The metadata should explicitly reference the dataset's globally unique and persistent identifier.
    - **Recommendation #3.3**. If no specific metadata standard is applied, at the very least, include the following fields:
        - Creator(s)
        - Creation date
        - Modified date
        - Publication date
        - Title
        - Identifier
        - Dataset description
        - Dataset language
        - Input data/parameters/filters/protocols used to generate the dataset
        - Ontology-aligned keywords and meta-tags
        - Version
        - Size
        - Format
        - Project acronym (OneSTOP) and number (101180559)
        - Licence of use
        - Embargo period

        - Publisher

- **Recommendation #4 - Indexing**. To ensure your research (meta)data are findable online, choose data repositories that register, index, and make your (meta)data discoverable.
- **Recommendation #5 - Access protocols**. Research data should be easily accessible online using standard, free and open protocols, rather than relying on specialised or proprietary tools or communication methods. If there are conditions and/or you are not using a free and open protocol for accessing the data, ensure that the requirements are clearly defined, allowing automated systems to handle or notify users.
- **Recommendation #6 - Language**. Use a formal, accessible, shared, and broadly applicable language for knowledge representation, such as the Resource Description Framework, Web Ontology Language, DARPA Agent Markup Language or the Javascript Object Notation for Linking Data.
- **Recommendation #7 - Standards and vocabularies**. Use data standards and vocabularies such as Darwin Core and the European nature information system (EUNIS) that are documented and accessible to anyone using the dataset.
- **Recommendation #8 - References**. Include qualified references to and properly cite other (meta)data by specifying if one dataset builds on another, if additional datasets are needed to complete the data, or if complementary information is stored in a different dataset.
- **Recommendation #9 - Clear licensing**. Release data with a clear data usage licence.
- **Recommendation #10 - Supporting documentation**. Information should be provided about any research output or any other tools and instruments needed to re-use or validate the data. Describe data not only with standardised metadata supporting discovery but also provide readme files or documentation needed for data analysis, validation and reuse with information on data provenance, scope, limitations, generation process, etc., such as the following:
    - Data origin;
    - Data provider;
    - Date of data retrieval;
    - Licence of use;
    - A URL or DOI for the original data;
    - Description of the input data, including its format, structure, and content;
    - Steps taken to clean and preprocess the data to ensure data quality and integrity.
- **Recommendation #11 - Open data rules**. According to Article 17 of the Grant Agreement, OneSTOP partners should make their data open access[18] by:
    - **Recommendation #11.1**. Uploading the data to an open access trusted data repository under the latest available version of the Creative Commons Attribution International Public License (CC BY) or Creative Commons Public Domain Dedication (CC0) or a licence with equivalent rights. Partners should preferably use a thematic repository like GBIF. If no such repository exists, they could use a generic repository such as Zenodo or a trusted institutional repository.

[18] Unless providing open access would, in particular, 1) be against the beneficiary's legitimate interests, including regarding commercial exploitation, or 2) be contrary to any other constraints, in particular the EU competitive interests or the beneficiary's obligations. If open access is not provided (to some or all data), partners must inform Pensoft in order to include a justification in the DMP.

    - **Recommendation #11.2**. The data underlying scientific publications should be deposited in a repository at the time of paper publication and made openly accessible as soon as possible, but no later than 12 months after the paper's publication. Other research data should be deposited and made open at the earliest opportunity.
    - **Recommendation #11.3**. Information should be provided via the repository about any research output or any other tools and instruments needed to re-use or validate the data.
    - **Recommendation #11.4**. The metadata of deposited data must put in the public domain under a CC0 licence waiver or another equivalent waiver (to the extent legitimate interests or constraints are safeguarded). Datasets may become unavailable over time due to maintenance costs, but metadata should persist, even if the data are no longer accessible. This ensures valuable information remains searchable and useful for research, including in replication studies.[19]
- **Recommendation #12 - Management of other research outputs**. OneSTOP partners are encouraged to manage their other research outputs in accordance with the FAIR principles described in Chapter 3.:
    - **Recommendation #12.1.** Strive to deposit your research outputs described with rich metadata in trusted open access repositories which assign them globally unique and persistent identifiers and offer search engines and indexing.
    - **Recommendation #12.2.** Use standard formats, vocabularies and ontologies and accompany your outputs with a separate human-readable description of the output, where needed.
    - **Recommendation #12.3.** Strive to make your outputs accessible under open licences, such as CC BY, CC0 and the MIT Licence. When using the MIT Licence, consider using the full text and disclaimer of the licence:

      MIT License

      Copyright (c) [year] [full name]

      Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

      The above copyright notice and this permission notice (including the next paragraph) shall be included in all copies or substantial portions of the Software.

      THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR

[19] In cases where data are closed but there are no compelling reasons that the related metadata should be, it is recommended that open access is provided to the metadata of the data, with CC0 or equivalent, while the dataset itself remains closed.

PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

- **Recommendation #13 - Scientific publications**. OneSTOP partners should adhere to the following open access rules for publications:
    - **Recommendation #13.1.** Deposit in a trusted repository, at the latest at the time of publication, a machine-readable electronic version of the published version or the final peer-reviewed manuscript accepted for publication;
    - **Recommendation #13.2.** Provide immediate open access via the repository, under the latest available version of the Creative Commons Attribution International Public License (CC BY) or a licence with equivalent rights;
        - For monographs and other long-text formats, the licence may exclude commercial uses and derivative works (e.g., CC BY-NC, CC BY-ND);
    - **Recommendation #13.3.** Give information via the repository about any research output or any other tools and instruments needed to validate the conclusions of the scientific publication;
    - **Recommendation #13.4.** Include a Data Availability Statement (even when there is no associated data) with information on where data can be accessed, preferably not referring readers to contact authors in order to access the data.

- **Recommendation #14 - Software development guidelines:**
    - **Recommendation #14.1.** Code should be developed in a public GitHub repository as part of the OneSTOP organisation. The repository should be set up in a structured way and contain a README.md, .gitignore, CITATION.cff, CODE_OF_CONDUCT.md and LICENCE file with mention of the copyright holder (recommended: MIT). Code that already exists (locally), should be added to GitHub following these guidelines. In case code has already been developed in a private GitHub repository, the repository should be transferred to the OneSTOP organisation (rights to transfer can be requested with one of the organization owners).
    - **Recommendation #14.2.** Code contributions should follow the GitHub flow. The repository's main branch should contain the software code in a state that can be installed without issue. Changes by collaborators should be made in separate branches and should be reviewed by collaborators through a pull request before merging into the main branch.
    - **Recommendation #14.3.** Developed code should be properly documented. Some guidelines for R function documentation can be found here. The repository should also contain a README file which explains the rationale and scope of the software, the instructions for installation, and a minimal example of the software workflow. Additional examples can be written out in the form of tutorials, which can be included on the project website.
    - **Recommendation #14.4.** Semantic versioning should be used throughout the software development process. Major and minor versions of the software should have an associated GitHub release, and should be published automatically to Zenodo. Version changes should be communicated and explained in a changelog in the form of a NEWS.md file which is added to the repository and kept up-to-date.

- **Recommendation #14.5.** Code should be written in open-source languages, and best practices for coding style should be followed (as detailed in the B-Cubed software development guide).
- **Recommendation #14.6.** Software developed within OneSTOP should be subjected to a review process. The maintainer of the software should reach out to OneSTOP partners, or to colleagues from their organisation, to execute a voluntary software review of their developed code, at least once and before the deliverable pertaining to software is submitted.

10. Annexe 1: Data Management Survey

0%

OneSTOP Data Management Questionnaire

To obtain a better understanding of why we collect this information, you can consult **this template (https://enspire.science/wp-content/uploads/2021/09/Horizon-Europe-Data-Management-Plan-Template.pdf)** provided by the European Commission.

We realise you probably don't have the exact answers to some of these questions yet. Feel free to give us your preliminary estimate, which you can modify/update during the mandatory DMP updates in M20 (August 2026) and M40 (April 2028).

In the DMP, we will clarify that the information it contains is based on an initial data mapping and is subject to change in order to reflect the project's development and arising data needs.

*1. First and last name

*2. Organisation/institution

Questions 3 to 5 concern the **data you will generate**.

3. Please provide the following *modifiable* information for the data you plan to **generate.** If you would like to add information for more than 5 datasets, please submit a second entry to the survey.

1. **Name of the data output**: this refers to the finalised data product, for example, a single dataset.

2. **Contact person**: the person who can be approached with questions regarding how the data is stored, for example.

3. **Relevant task**: task in which the dataset will be generated.

4. **Generated via**: for example, observation data, data aggregation, modelling, remote sensing, literature review, policy review, interview, surveys.

5. **Size**: a rough estimate which can be revised/updated when significant changes occur.

6. **Format**: for example, .csv; .netcdf .docx; .xlsx; .pdf; .mp4; .xml.

7. **Sensitive data**: Yes/no. If yes, please specify, for example, racial, political, ethical, health, and more here (https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/legal-grounds-processing-data/sensitive-data/what-personal-data-considered-sensitive_en).

8. **Personal data**: Yes/no. If yes, please specify, for example, name, surname, address, email, IP address, location data.

9. **Species of conservation concern**: Does the dataset contain observation data on species of conservation concern that might lead to their exploitation if their sites are not protected?

10. **Publication timeline**: a rough estimate of when the data can be publicly shared (e.g., M20). If there is an embargo period, please specify why and how long it will apply.

11. **Metadata**: will you use common metadata standards (e.g., EML, INSPIRE, Darwin Core)? If not, what fields will you include in your metadata? For example, DOI, creator(s), title(s), date(s), size(s), format(s), rights, description(s), included datasets, applied filters.

12. **Users**: to whom might this data be useful?

13. **Type of access**: will it be open access? If not, please indicate the reasons, for example, ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related, contract.

14. **Licence**: what type of licence do you plan to use for your data? Keep in mind, in the Grant Agreement the following was promised:

Datasets of species observations and checklists will be made available under a Creative Commons Zero waiver (CC0). Other datasets will also be made available either in the public domain or with an attribution licence depending on the nature of the material and the wishes of the owners.

15. **Point of access**: do you have a preferred route for making this data open access? For example, Zenodo, GBIF, GitHub, etc.

16. **Re-use**: potential documents or tools needed to re-use or validate the data.

❓ This question concerns the data you will **generate**.

❗ **Please fill in at least one answer**

	Data output 1	Data output 2	Data output 3	Data output 4	Data output 5
Name of the data output					
Contact person					
Relevant task					
Generated via					

	Data output 1	Data output 2	Data output 3	Data output 4	Data output 5
Size					
Format					
Sensitive data					
Personal data					
Species of conserva-tion concern					
Publication timeline					
Type of metadata					
Potential users					
Type of access					
Licence					
Point of access					

	Data output 1	Data output 2	Data output 3	Data output 4	Data output 5
Re-use					

*4. If your **generated data** contains personal or sensitive information, would you consider anonymising it and publishing it in an aggregated form so as to not disclose said information?

If yes, please specify to which dataset(s) that would apply. If not, please use the box to explain why.

This question concerns the data you will **generate**.

Choose one of the following answers

- Yes
- No
- To be determined
- Not relevant

Please enter your comment here:

*5. Will you generate digital sequence information derived from genetic resources? If yes, please specify in the box:

5.1 Are these intended to be made publicly available?

5.2 Do these need to be governed by mutually agreed terms?

5.3 Are these data covered by other international access and benefit sharing agreements?

Choose one of the following answers

- ○ Yes
- ○ No

Please enter your comment here:

If you have any additional information/comments related to questions 3-5, please provide them in this field.

Question 6 concerns the **data you will obtain from elsewhere and reuse**.

6. Please provide the following *modifiable* information for the **existing data you will reuse.** If you would like to add information for more than 5 datasets, please submit a second entry to the survey.

1. **Data name**
2. **Relevant task**
3. **Size**
4. **Format**: for example, .docx; .xlsx; .pdf; .mp4; .xml; .csv.

5. **Sensitive data**: Yes/no. If yes, please specify, for example, racial, political, ethical, health, and more here (https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/legal-grounds-processing-data/sensitive-data/what-personal-data-considered-sensitive_en).

6. **Personal data**: Yes/no. If yes, please specify, for example, name, surname, address, email, IP address, location data.

7. **Metadata**: describe the metadata accompanying the data.

8. **Access**: open/restricted/closed access.

9. **Origin**: what is the origin of the data?

10. **Ownership**: who owns the data you will reuse?

11. **Licence**: what type of licence allows you to use the data?

12. **Re-use**: potential documents or tools needed to re-use or validate the data.

This question concerns the data you will obtain from elsewhere and **reuse**.

	Data 1	Data 2	Data 3	Data 4	Data 5
Data name					
Relevant task					
Size					
Format					
Sensitive data					
Personal data					

	Data 1	Data 2	Data 3	Data 4	Data 5
Metadata					
Access					
Origin					
Ownership					
Licence					
Re-use					

If you have any additional information/comments related to question 6, please provide them in this field.

According to the European Commission, in addition to the management of data, beneficiaries should also consider and plan for the management of other research outputs that may be generated or re-used throughout their projects. Such outputs can be either digital (e.g. software, workflows, training materials, protocols, models, etc.) or physical (e.g. hardware, new materials, antibodies, reagents, samples, etc.). Beneficiaries should strive to provide sufficient detail on how their research outputs will be managed and shared, or made available for re-

use, in line with the FAIR principles.

Question 7 concerns the **other research outputs you will generate** (e.g. software, workflows, training materials, models, hardware, etc.).

7. Please provide the following *modifiable* information for the **other research outputs you plan to generate.** If you would like to add information for more than 5 outputs, please submit a second entry to the survey.

1. **Name of the output**

2. **Contact person**: the person who can be approached with questions regarding the development and publication of the output, for example.

3. **Relevant task**: task in which the output was produced.

4. **Type of output**: briefly describe the result you will produce.

5. **Delivery timeline**: a rough estimate of when the output could be publicly shared.

6. **Users**: to whom might this output be useful?

7. **Metadata**: will you accompany the publication of your result with metadata? If yes, will you use a common metadata format and/or what fields will you include in your metadata?

8. **Type of access**: will it be open access? If not, please indicate the reasons, for example, intellectual property, commercial, privacy-related, security-related.

9. **Persistent identifiers**: do you plan on sharing your output in a way that assigns it globally unique and persistent identifiers?

10. **Point of access**: do you have a preferred route for making your output accessible to end users?

11. **Interoperability**: how do you plan to make your output interoperable?

12. **Licence**: what type of licence do you plan to use for your output?

13. **Intellectual property**: Please shortly describe the expected intellectual property management of this output. Who will own the IP rights? Are there any existing patents, trademarks, or copyrights related to this product? If not, what type of intellectual property protection would apply to the output (N/A, copyright, patent, trademark, other)?

14. **Re-use**: potential documents or tools needed to re-use the output.

❓ This question concerns the other research outputs you will **generate**.

❗ **Please fill in at least one answer**

	Output 1	Output 2	Output 3	Output 4	Output 5
Name of the output					
Contact person					
Relevant task					
Type of output					
Delivery timeline					
Users					
Metadata					
Type of access					
Identifier					
Point of access					

	Output 1	Output 2	Output 3	Output 4	Output 5
Interoperability					
Licence					
Intellectual property					
Re-use					

If you have any additional information/comments related to question 7, please provide them in this field.

Question 8 concerns the **other research outputs you will reuse** (e.g. software, workflows, training materials, models, hardware, etc.).

8. Please provide the following *modifiable* information for the **other research outputs you plan to reuse**:

1. **Name of the output**
2. **Relevant task**

3. **Access**: open/restricted/closed access.

4. **Origin**: what is the origin of the output?

5. **Ownership**: who owns the output you will reuse?

6. **Licence**: what type of licence allows you to use the output?

7. **Re-use**: potential documents or tools needed to re-use the output.

This question concerns the other research outputs you will obtain from elsewhere and **reuse**.

	Output 1	Output 2	Output 3	Output 4	Output 5
Name of the output					
Relevant task					
Access					
Origin					
Ownership					
Licence					
Re-use					

If you have any additional information/comments related to question 8, please provide them in this field.

Questions 9-13 concern your **data management practices**.

*9. Please provide a brief summary of your **institutional data management practices**, specifying:

1. **Data location**: where and how data will be stored before it is fully compiled, for example, institutional server or web hosting.

2. **Server location**: EU or non-EU; compliant or non-compliant with applicable data protection rules (for example, GDPR).

3. **Backup procedures**: type of backup procedures and their frequency.

4. **Protection**: how data security is ensured, for example, password or two-factor authentication.

5. **Data Protection Officer (DPO)**: name the person from your organisation who is responsible for data protection and can serve as a contact person if questions arise.

This question concerns your **institutional data management**.

Data location	
Server location	
Backup procedures	

Protection	
Data Protection Officer (DPO)	

✱10. Will you be using standard vocabularies (e.g., Dublin Core) for all data types present in your data to allow inter-disciplinary interoperability?

If yes, please use the box to specify.

❓ This question concerns your **data management**.

❗ Choose one of the following answers

- ○ Yes
- ○ No

Please enter your comment here:

✱11. Do you follow a specific naming convention for your files?

If yes, please use the box to specify.

For example: [OneSTOP_filename_version_creation-date], where data format should be YYYYMMDD, number style version should be 01, 02, 03.

❓ This question concerns your **data management**.

❗ Choose one of the following answers

- Yes
- No

Please enter your comment here:

✱12. Do you have a preference for a trusted repository where to publish your public research data?

If yes, please use the box to specify.

This question concerns your **data management**.

Choose one of the following answers

- Yes
- No

Please enter your comment here:

✱13. Can you identify potential obstacles (e.g., technical, social, policies) that would prevent delivering FAIR data during OneSTOP's lifetime and beyond? Information on FAIR data here (https://www.openaire.eu/how-to-make-your-data-fair).

If yes, please use the box to specify.

This question concerns your **data management**.

Choose one of the following answers

- Yes
- No

Please enter your comment here:

If you have any additional information/comments related to questions 9-13, please provide them in this field.

Questions 14 and 15 concern **ethics**.

✱14. Does your research activity involve/fall within any of the following identified '**four critical areas of ethical concern**'?

If yes, please provide a short statement describing your work and how you will follow OneSTOP's ethics self-assessment statement (page 145 of the Grant Agreement), the European Code of Conduct for Research Integrity (https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/european-code-of-conduct-for-research-integrity_horizon_en.pdf) and GDPR (https://gdpr-info.eu/).

❗ Comment only when you choose an answer.

☐ (a) Involvement of humans in the project (as volunteer research participants)	
☐ (b) Participation of non EU countries	
☐ (c) Research on Fauna, Flora and Protected Areas	
☐ (d) Utilisation of Artificial intellegence	
☐ Not applicable	

✱15. Which option best describes the current state of the **ethical approval** for your work in OneSTOP?

For answers (c),(d), or (e), please provide further information in the box. For example, when will ethics approval be secured; why are you unable to secure it; why is it not relevant?

❗ Choose one of the following answers

- ○ (a) I already secured ethical approval for my research
- ○ (b) I submitted but not yet secured ethical approval for my research
- ○ (c) I have not yet secured ethics approval for my research
- ○ (d) I am unable to secure ethics approval
- ○ (e) This is not relevant for my research

Please enter your comment here:

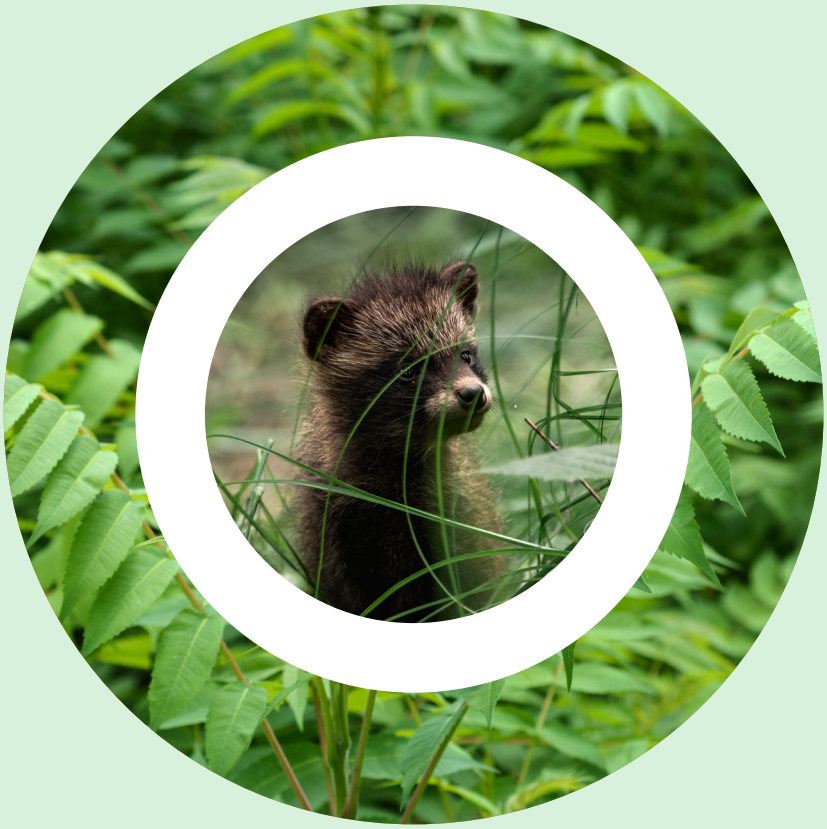
Please do not hesitate to fill in the comments fields with any information that would help to clarify your answer or is not addressed in the list of questions!

11. Annexe 2: Data Management Guidelines

OneSTOP

onestop-project.eu

OneBiosecurity Systems and Technology for People, Places and Pathways

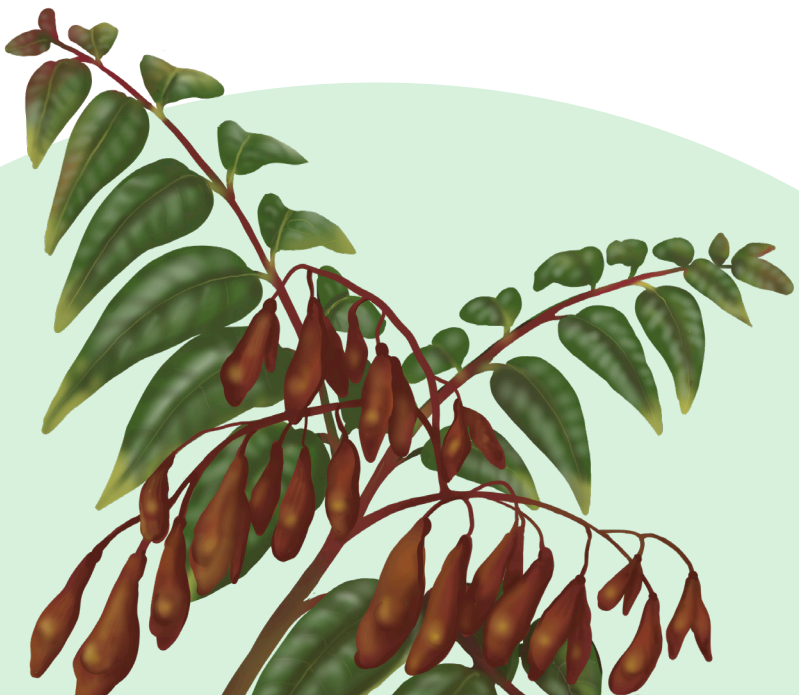


Data Management Guidelines

Funded by the European Union

OneSTOP receives funding from the European Union's Horizon Europe Research and Innovation Programme (ID No 101180559). Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the EU nor REA can be held responsible for them.

TABLE OF CONTENTS



RECOMMENDATION 1 | TRUSTED REPOSITORIES

Upload datasets to trusted repositories[1] and infrastructures such as Zenodo or GBIF, which assign a DOI and allow easy download via a simple https link, without the need for special tools for access. This allows anyone with the link to retrieve the data through a standard web browser, provided the data are open. In terms of a general-purpose trusted repository, the majority of OneSTOP partners expressed a preference for Zenodo as a suitable option, which they are familiar with and have previously explored appropriate arrangements regarding deposition. Thus, when not using domain-specific repositories, OneSTOP partners are encouraged to deposit their datasets to Zenodo as the project can verify that it covers all the **requirements of FAIR data**, most importantly:

- A DOI is issued to every published dataset on Zenodo;
- Metadata of each record are indexed and searchable directly in Zenodo's search engine immediately after publishing;
- Metadata of each record are sent to DataCite servers during DOI registration and indexed there;
- Metadata for individual records as well as record collections are harvestable using the standard, open, free and universal OAI-PMH protocol by the record identifier and the collection name;
- Metadata are publicly accessible and in the public domain. No authorisation is ever required to retrieve it;
- Data and metadata will be retained for the lifetime of the repository, ensuring that the metadata will be accessible, even when the data are no longer available;
- Metadata are stored in high-availability database servers at CERN, which are separate from the data itself.

[1] Trusted repositories are either certified repositories (e.g., CoreTrustSeal) and/or disciplinary/domain repositories that are commonly used/endorsed by research communities (e.g., ELIXIR deposition databases). General-purpose (e.g., Zenodo) and institutional repositories could also be acceptable if they share these properties: Mechanisms to ensure integrity and authenticity of contents; Clear information about their policies/services; Broad, ideally open access to content (consistent with legal and ethical constraints); Assigned PIDs; Detailed, machine-readable and standardised metadata; Mid- and long-term preservation of contents, expert curation, quality assurance; National and/or international security criteria.

RECOMMENDATION 2 | IDENTIFIERS

Publish data with globally unique and persistent identifiers. You can find a list of services that supply such identifiers **here**.

RECOMMENDATION 3 | METADATA

Provide machine-readable, detailed metadata adhering to standardised metadata formats. Include descriptive information about the context, quality and condition of the data, using keywords which enable computers to automatically sort and prioritise data.

RECOMMENDATION 3.1

OneSTOP has identified a number of suitable standards that can be used based on each partner's preference and specific data needs:

- **Ecological Metadata Language (EML)**: includes modules for the spatial, temporal, taxonomic, and thematic extent of data, as well as research methods and protocols.
- **DataCite Metadata**: provides a standardised framework for data across disciplines.

RECOMMENDATION 3.2

If no specific metadata standard is applied, at the very least, include the following fields:

- Creator(s)
- Creation date
- Modified date
- Publication date
- Title
- Identifier
- Dataset description
- Dataset language
- Input data/parameters/filters/protocols used to generate the dataset
- Ontology-aligned keywords and meta-tags
- Version
- Size
- Format
- Project acronym (OneSTOP) and number (101180559)
- Licence of use
- Embargo period
- Publisher

RECOMMENDATION 4 | INDEXING

To ensure your research (meta)data are findable online, choose data repositories that register, index, and make your (meta)data discoverable.

RECOMMENDATION 5 | ACCESS PROTOCOLS

Research data should be easily accessible online using standard, free and open protocols, rather than relying on specialised or proprietary tools or communication methods. If there are conditions and/or you are not using a free and open protocol for accessing the data, ensure that the requirements are clearly defined, allowing automated systems to handle or notify users.

RECOMMENDATION 6 | LANGUAGE

Use a formal, accessible, shared, and broadly applicable language for knowledge representation, such as the Resource Description Framework, Web Ontology Language, DARPA Agent Markup Language or the Javascript Object Notation for Linking Data.

RECOMMENDATION 7 | STANDARDS AND VOCABULARIES

Use data standards and vocabularies such as Darwin Core and the European Nature Information System (EUNIS) that are documented and accessible to anyone using the dataset.

RECOMMENDATION 8 | REFERENCES

Include qualified references to and properly cite other (meta)data by specifying if one dataset builds on another, if additional datasets are needed to complete the data, or if complementary information is stored in a different dataset.

RECOMMENDATION 9 | CLEAR LICENSING

Release data with a clear data usage licence.

RECOMMENDATION 10 | SUPPORTING DOCUMENTATION

Information should be provided about any research output or any other tools and instruments needed to re-use or validate the data. Describe data not only with standardised metadata supporting discovery but also provide readme files or documentation needed for data analysis, validation and reuse with information on data provenance, scope, limitations, generation process, etc., such as the following:

- Data origin;
- Data provider;
- Date of data retrieval;
- Licence of use;
- A URL or DOI for the original data;
- Description of the input data, including its format, structure, and content;
- Steps taken to clean and preprocess the data to ensure data quality and integrity.

RECOMMENDATION 11 | OPEN DATA RULES

According to Article 17 of the Grant Agreement, OneSTOP partners should make their data open access[2] by:

RECOMMENDATION 11.1

Uploading the data to an open access trusted data repository under the latest available version of the Creative Commons Attribution International Public License (CC BY) or Creative Commons Public Domain Dedication (CC0) or a licence with equivalent rights. Partners should preferably use a thematic repository like GBIF. If no such repository exists, they could use a generic repository such as Zenodo or a trusted institutional repository.

RECOMMENDATION 11.2

The data underlying scientific publications should be deposited in a repository at the time of paper publication and made openly accessible as soon as possible, but no later than 12 months after the paper's publication. Other research data should be deposited and made open at the earliest opportunity.

RECOMMENDATION 11.3

Information should be provided via the repository about any research output or any other tools and instruments needed to re-use or validate the data.

RECOMMENDATION 11.4

The metadata of deposited data must be open under CC0 or an equivalent licence (to the extent legitimate interests or constraints are safeguarded). Datasets may become unavailable over time due to maintenance costs, but metadata should persist, even if the data are no longer accessible. This ensures valuable information remains searchable and useful for research, including in replication studies[3].

[2] Unless providing open access would, in particular, 1) be against the beneficiary's legitimate interests, including regarding commercial exploitation, or 2) be contrary to any other constraints, in particular the EU competitive interests or the beneficiary's obligations. If open access is not provided (to some or all data), partners must inform Pensoft in order to include a justification in the DMP.

[3] In cases where data are closed but there are no compelling reasons that the related metadata should be, it is recommended that open access is provided to the metadata of the data, with CC0 or equivalent, while the dataset itself remains closed.

RECOMMENDATION 12 | MANAGEMENT OF OTHER RESEARCH OUTPUTS

OneSTOP partners are encouraged to manage their other research outputs in accordance with the FAIR principles described in Chapter 3 of the DMP::

RECOMMENDATION 12.1

Strive to deposit your research outputs described with rich metadata in trusted open access repositories which assign them globally unique and persistent identifiers and offer search engines and indexing.

RECOMMENDATION 12.2

Use standard formats, vocabularies and ontologies and accompany your outputs with a separate human-readable description of the output, where needed.

RECOMMENDATION 12.3

Strive to make your outputs accessible under open licences, such as CC BY, CC0 and the MIT Licence. When using the MIT Licence, consider using the full text and disclaimer of the licence:

MIT License

Copyright (c) [year] [full name]

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice (including the next paragraph) shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

RECOMMENDATION 13 | SCIENTIFIC PUBLICATIONS

OneSTOP partners should adhere to the following open access rules for publications:

RECOMMENDATION 13.1

Deposit in a trusted repository, at the latest at the time of publication, a machine-readable electronic version of the published version or the final peer-reviewed manuscript accepted for publication;

RECOMMENDATION 13.2

Provide immediate open access via the repository, under the latest available version of the Creative Commons Attribution International Public Licence (CC BY) or a licence with equivalent rights;

For monographs and other long-text formats, the licence may exclude commercial uses and derivative works (e.g., CC BY-NC, CC BY-ND);

RECOMMENDATION 13.3

Give information via the repository about any research output or any other tools and instruments needed to validate the conclusions of the scientific publication;

RECOMMENDATION 13.4

Include a Data Availability Statement (even when there is no associated data) with information on where data can be accessed, preferably not referring readers to contact authors in order to access the data.

RECOMMENDATION 14 | SOFTWARE DEVELOPMENT GUIDELINES

RECOMMENDATION 14.1

Code should be developed in a public GitHub repository as part of the **OneSTOP organisation**. The repository should be set up in a structured way and contain a README.md, .gitignore, CITATION.cff, CODE_OF_CONDUCT.md and LICENSE file with mention of the copyright holder (recommended: MIT). Code that already exists (locally), should be added to GitHub following **these guidelines**. In case code has already been developed in a private GitHub repository, the repository should be transferred to the OneSTOP organisation (rights to transfer can be requested with one of the organization owners).

RECOMMENDATION 14.2

Code contributions should follow the **GitHub flow**. The repository's main branch should contain the software code in a state that can be installed without issue. Changes by collaborators should be made in separate branches and should be reviewed by collaborators through a pull request before merging into the main branch.

RECOMMENDATION 14.3

Developed code should be properly documented. Some guidelines for R function documentation can be found **here**. The repository should also contain a README file which explains the rationale and scope of the software, the instructions for installation, and a minimal example of the software workflow. Additional examples can be written out in the form of tutorials, which can be included on the project website.

RECOMMENDATION 14.4

Semantic versioning should be used throughout the software development process. Major and minor versions of the software should have an associated **GitHub release**, and should be published automatically to Zenodo. Version changes should be communicated and explained in a changelog in the form of a NEWS.md file which is added to the repository and kept up-to-date.

RECOMMENDATION 14.5

Code should be written in open-source languages, and best practices for coding style should be followed (as detailed in the B-Cubed **software development guide**).

RECOMMENDATION 14.6

Software developed within OneSTOP should be subjected to a review process. The maintainer of the software should reach out to OneSTOP partners, or to colleagues from their organisation, to execute a voluntary software review of their developed code, at least once and before the deliverable pertaining to software is submitted.